

BEFORE THE  
POSTAL REGULATORY COMMISSION  
WASHINGTON, D.C. 20268-0001

PERIODIC REPORTING  
(PROPOSAL THIRTEEN)

Docket No. RM2015-7

**RESPONSES OF THE UNITED STATES POSTAL SERVICE  
TO QUESTIONS 1-16 AND 19-28 OF CHAIRMAN'S  
INFORMATION REQUEST NO. 1  
(January 12 , 2015)**

The United States Postal Service hereby provides its responses to Questions 1-16 and 19-28 of Chairman's Information Request No. 1, issued January 6, 2015. The questions are stated verbatim and followed by the response. Responses to Questions 17 and 18 are still being prepared.

Respectfully submitted,

UNITED STATES POSTAL SERVICE

By its attorney:

---

Eric P. Koetting

475 L'Enfant Plaza West, S.W.  
Washington, D.C. 20260-1137  
(202) 277-6333  
January 12, 2015

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

1. Please refer to USPS-RM2015-7/1 "Report on the City Carrier Street Time Study" (Report). On page 76, the Postal Service states: "Another issue that bears investigation is the relatively high marginal time for FSS mail.... The normal way to proceed with this investigation would be to separately estimate the regular delivery equation for FSS and non-FSS ZIP Code days. However, there is a problem with this approach because of the relatively small number of FSS ZIP Code days. With just 967 observations for FSS zones, multicollinearity becomes a serious problem. Even with the reduced-variable model, twelve of twenty-seven estimated coefficients have low t-statistics in an FSS-only model."
  - a. Please provide the results of the separately estimated delivery equation for FSS and non-FSS ZIP Code days, and the t-statistics and other diagnostic results used to evaluate it.
  - b. Please discuss how the low t-statistics of the separately estimated equation for FSS and non-FSS ZIP Code days limit its use as compared to the broader model approach. Specifically, please detail what t-statistic results would need to be achieved for the benefits of this approach to outweigh the concerns about the level of significance of the variables.

**RESPONSE:**

- a. Please see the table on the following page.
- b. As that table shows, twelve of the coefficients to be estimated have t-statistics so low that one cannot reject the null hypothesis that the coefficient is equal to zero. This has serious implications for the estimation. First, it shows that multicollinearity is a material problem. Second, it greatly reduces the usefulness of the model. For example, all of the direct coefficients for both cased mail and volumes collected are estimated to be zero, seriously calling into question the model's ability to accurately estimate a variability for those two shapes. For this approach to be sufficiently reliable, one would want to see results for the t-

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

statistics that are comparable to those in either the original equation without the FSS dummy variable, or the equation that includes the FSS dummy variable.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

Results of Estimation of the Regular Delivery Equation  
PRC Reduction Method-FSS Zones Only

<b>Variable</b>	<b>Coefficient</b>	<b>t-statistic</b>	<b>H.C. t-statistic</b>
Intercept	-29.13	-10.23	-12.29
DPS	3.07	4.81	4.42
DPS2	-0.00002	-3.88	-3.98
CM	1.03	0.49	0.56
CM2	-0.00004	-0.59	-0.58
SEQ	2.77	5.06	4.47
SEQ2	-0.00005	-3.16	-2.60
FSS	4.50	2.36	2.43
FSS2	-0.00008	-1.44	-1.53
CV	-1.19	-0.33	-0.33
CV2	0.00005	0.25	0.22
DP	28.33	13.88	14.03
DP2	-0.00064	-8.07	-8.84
DPS*CM	0.00003	0.76	0.85
DPS*CV	-0.00013	-2.17	-2.06
DPS*PD	0.00008	1.97	1.81
CM*CV	0.00069	3.68	3.59
CM*DP	-0.00008	-0.66	-0.65
FSS*CV	-0.00002	-0.12	-0.12
FSS*DP	0.00006	0.67	0.63
CV*DP	0.00038	2.27	2.23
DM	54.14	7.75	8.43
DM2	-33.62	-4.98	-5.34
MPDP	6358.90	3.41	3.76
MDPD2	-946704.00	-2.89	-3.64
BR	56.70	0.97	1.11
BR2	-314.91	-1.37	-1.8

*Cost driver coefficients are in seconds; characteristic variable coefficients are in hours.*

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

2. On page 34 of the Report, the Postal Service states: "Prudence dictated that imputation was considered only for ZIP Code days for which at least 80 percent of the routes reported volume."
- a. Please describe how this 80 percent threshold was determined.
  - b. Please detail the impact of this choice of threshold.

**RESPONSE:**

a. The 80 percent threshold was determined from a review of ZIP Codes that had missing data for some routes. The 80 percent threshold provided a demarcation between those ZIP Codes that had just a few routes with missing data, and the ZIP codes that had multiple routes with missing data. The former are reasonable candidates for imputation because nearly all of the data are in place. The ZIP Codes with a high proportion of routes missing did not meet this criterion.

b. Conceptually, the impact of this threshold is to identify those ZIP Codes that have small proportions of routes with missing data and are thus candidates for imputation. Moreover, variations in the value used for the threshold could affect the number of imputations made. If the threshold value was higher, then fewer route days would be considered for imputation. In contrast, if the threshold value was lower, more route days would be considered for imputation. The important point is that, quantitatively, variations in the value for the threshold are likely to have a *de minimis* impact on the results. This is because so few imputations were done. Only 206 of over 72,000 route days had imputed data. Variations in

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

the threshold value would not have a material impact on the number of route days in the data set. For example, if the threshold were raised to 100 percent, so no imputation was done, the number of routes days in the data set would change only by a tiny amount.

Moreover, when imputation was done, it was often done for just one route in a ZIP Code, so the imputation likely had relatively small impact on the ZIP Code's daily collection volume. Nevertheless, imputation permitted the preservation of data and provided the advantage of ensuring that all ZIP Code days used in the analysis were complete.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

3. Please discuss how Segment 7 cost pool data will be updated in subsequent fiscal years. If route evaluations will be used to form Segment 7 cost pools, will they be formed using the same criteria discussed in Section D of the Report (pages 9-14)?

**RESPONSE:**

Cost pool proportions will be updated by using the Route Evaluation System data. A data extract will be pulled from that system and, as described in the Report, that data extract will be used to calculate the proportions. The formation of cost pools will, in general, follow the same criteria discussed in Section D of the Report, but a change in circumstances could necessitate review of those criteria.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

4. Please refer to page 10 of the Report, where the Postal Service states that there were 37 route evaluations reporting gross street time over 12 hours and that since this outcome is not possible, these evaluations were dropped. Were these 37 routes examined to determine if any of them included pivot time from another route, which could suggest that more than 12 hours may have been a possible outcome?

**RESPONSE:**

While pivot time from other routes could cause the recorded street time on a given route to exceed eight hours, discussions with operation experts indicated that it would be extremely rare to have over twelve hours of street time. It is far more likely that 12 hours of street time reflected either an error in data collection or that the route day was subject to an extraordinary circumstance.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

5. Please refer to the excel files contained in the USPS-RM2015-7/1, WEBTOOL\_Excel\_Files\_Parcel\_Accountable\_Study folder, and the variables listed in USPS-RM2015-7/1, Deviation\_Acct\_Variabilities\_Model.SAS and the In\_Receptacle\_Variabilities\_Model.SAS.
- a. Please define and identify the mail contained in the following column headings or variables:
    - i. PP\_EXPRESS
    - ii. PP\_PRIORITY
    - iii. PP\_FC
    - iv. PP\_OTHER
    - v. ON\_DEMAND\_PICKUP\_STOPS
    - vi. ON\_DEMAND\_STOPS\_MADE
    - vii. ON\_DEMAND\_EXPRESS ON\_DEMAND\_PRIORITY
    - viii. ON\_DEMAND\_FC
    - ix. ON\_DEMAND\_OTHER
    - x. ACCOUNTABLE DEVIATION IN\_RECEPTACLE
    - xi. ON\_DEMAND\_EXPRESS
    - xii. ON\_DEMAND\_FC
    - xiii. ON\_DEMAND\_OTHER
    - xiv. ON\_DEMAND\_PICKUP\_STOPS
    - xv. ON\_DEMAND\_PRIORITY
    - xvi. ON\_DEMAND\_STOPS\_MADE
    - xvii. PP\_EXPRESS PP\_FC PP\_OTHER
    - xviii. PP\_PRIORITY PP\_STOPS\_MADE
  - b. Please identify which of the variables listed above are carrier pickup items.
  - c. Please confirm that carrier pickup mail or time associated with this mail are not included in the estimation of variabilities for the regular delivery model, the parcel/accountable model or the in-receptacle model.
  - d. Please explain what role, if any, the variables listed above played in these models.
  - e. Please identify the cost pools to which each of the variables listed above are assigned.

**RESPONSE:**

- a. The following table lists and defines the variables requested. All of these variables were captured by route day.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

Variable	Definition
PP_EXPRESS	Volume of Express (now renamed Priority Mail Express) collected via Package (formerly Carrier) Pickup.
PP_PRIORITY	Volume of Priority Mail collected via Package Pickup.
PP_FC	Volume of First-Class Mail collected via Package Pickup.
PP_OTHER	Volume that was not Priority Mail Express, Priority, or First-Class collected via Package Pickup.
PP_STOPS_MADE	Number of stops made due to Package Pickup requests.
ON_DEMAND_PICKUP_STOPS	Number of <i>scheduled</i> On Demand pickup stops.
ON_DEMAND_STOPS_MADE	Number of <i>actual</i> On Demand pickup stops.
ON_DEMAND_EXPRESS	Volume of Priority Mail Express collected via On Demand pickup.
ON_DEMAND_PRIORITY	Volume of Priority Mail collected via On Demand pickup.
ON_DEMAND_FC	Volume of First-Class Mail collected via On Demand pickup.
ON_DEMAND_OTHER	Volume that was not Priority Mail Express, Priority Mail, or First Class collected via On-Demand pickup.
ACCOUNTABLE	Volume of accountables (those that require customer contact) to be delivered.
DEVIATION	Volume of Deviation (parcels which <i>do not</i> fit in the customer's letters and flats receptacle) parcels to be delivered.
IN_RECEPTACLE	Volume of In-receptacle (parcels that <i>do</i> fit in the customer's letters and flats receptacle) parcels to be delivered.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

b. The variables that begin with the characters 'PP' refer to either volumes or stops collected via Package (formerly Carrier) Pickup. The specific list of variables that refer to Package Pickup is 1. PP\_EXPRESS, 2. PP\_PRIORITY, 3. PP\_FC, 4. PP\_OTHER, and 5. PP\_STOPS\_MADE.

c. Confirmed. The volumes for carrier pickup were miniscule, so it wasn't feasible to include them in the equations. The median route day volumes were zero for all four products: PP\_EXPRESS, PP\_PRIORITY, PP\_FC, and PP\_OTHER. In addition, as the following table shows, almost no route days had any positive carrier pickup volumes.

Percentage of Route Days with No Carrier Pickup  
Volume

Product	Percentage
EXPRESS	98.6%
FIRST CLASS	93.4%
PRIORITY	90.3%
OTHER	96.6%

d. The variables, ACCOUNTABLE and DEVIATION, were volume cost drivers in the deviation delivery time equation. The variable IN\_RECEPTACLE was a volume cost driver in the in-receptacle delivery time equation.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

e. Please note that these are volume variables, not time variables, so none of them would be assigned to cost pools. Rather, they could serve as cost drivers for certain cost pools. As discussed in parts c. and d., the only variables which served as cost drivers would be ACCOUNTABLE, DEVIATION and IN\_RECEPTACLE. To the extent the question is asking where the time for carrier pickup or on-demand pickup would be in the city carrier cost model, the answer is that it would be in indirectly attributable time.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

6. Please refer to the Report, Table 6 on page 14, rows entitled "Parcel & Accountable Delivery" and "Relay."
- a. Please confirm that the Parcel & Accountable Delivery cost pool as a share of the total increases at a faster rate than Relay between FY 2009 and FY 2011. If not confirmed, please explain.
  - b. Please confirm that the Parcel/Accountable Delivery cost pool percent continues an upward trend from FY 2011 to FY 2013, while Relay declines during this period. If not confirmed, please explain.
  - c. Please confirm that the Network Travel cost pool has continuously declined as a share of the total between FY 2009 and FY 2011. If not confirmed, please explain.
  - d. Was a reduction in relay time or network travel time a goal of route reconfiguration, or were these reductions unintended consequences?

**RESPONSE:**

- a. Confirmed.
- b. Confirmed.
- c. Confirmed
- d. The reductions in the proportions of both relay time and network travel time reflect a Postal Service effort to reduce street time through developing more efficient routes. The increase in the proportion for parcel/accountable delivery time reflects the growth in package volume.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

7. Please refer to the Report, Table 7, page 17.
- a. Please confirm that the Column entitled "Ratio of Total Street Time to Attributable Street Time" is the ratio of (Direct + Indirect Street Time)/(Direct Street Time). If not, please explain.
  - b. Please confirm that this ratio is applied to both in-receptacle and deviation parcels. If not, please explain.

**RESPONSE:**

- a. The operations definition of street time includes vehicle load and unload time.

Because vehicle load and unload time is considered office time in the city carrier cost model, the relevant ratio is given by:

$$\frac{(\text{Directly Attributable Street Time} + \text{Indirectly Attributable Street Time} + \text{Vehicle Load and Unload Time})}{\text{Directly Attributable Street Time}}$$

- b. Confirmed that the ratio defined above is applied to both deviation and in-receptacle packages.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

8. Please refer to the Report, Table 9, page 19.
- a. Please confirm that cost pool shares for the activities listed in the table are based on direct plus allied street time shares, adjusted for in-receptacle parcels. If not, please explain.
  - b. Please explain whether the Segment 7 cost pool shares and/or the street time used to estimate shape variabilities in R2005-1 were based on direct street time for the activities listed in Table 9.

**RESPONSE:**

a. Not confirmed. The cost pool shares in Table 9 are the relative proportions of directly attributable street time, adjusted for in-receptacle package delivery time.

As Figure 2 on page 8 of the report shows, some parts of allied time are in indirectly attributable street time, and thus are not part of the cost pool shares in Table 9.

b. The cost pool shares and the times used to estimate the shape variabilities in the analysis presented in Docket No. R2005-1 were based upon directly attributable street time, just as they are in the instant analysis.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

9. This question seeks information on the formation of three SAS datasets in USPS-RM2015-7/1: package\_study\_volume\_masked\_zips.sas7bdat, scanrouteday\_masked\_zips.sas7bdat, and carrier\_scans\_pa\_study.sas7bdat.
- a. Please identify which files were created from raw scan data imported using SAS or another program, such as excel.
  - b. If the (SAS) files created from raw data were modified to create the datasets listed above, please provide electronic copies of the raw data, the SAS program(s) which created the final SAS dataset created from raw scan data, and the associated SAS logs. Please also explain each of the initial variables read in and describe the process of creating the final dataset.
  - c. Please provide a description of the variables EVENT\_DATE1, EVENT\_TIME1, EVENT\_CODE, TRACK\_BARCODE\_ID, TRDV\_DEV\_ID, barcode2, and barcode1.
  - d. Please confirm there were 19 Event Codes: 01, 02, 03, 04, 05, 06, 12, 13, 14, 15, 23, 45, 51, 52, 53, 54, 55, 56, and 59.
    - i. If confirmed, please explain each component or action of the event, from start to finish, for each of the 19 event codes.
    - ii. In not confirmed, please list all event codes and explain the nature of each code.
  - e. Please confirm that barcode1 is the beginning of the activity associated with an event code and barcode2 is the end of the activity associated with that event code. If not, please explain.

**RESPONSE:**

- a. The raw scan data is included in the SAS data set named Carrier\_scans\_pa\_study.sas7bdat. In organizing the raw data, a small number of duplicate and extraneous observations were dropped. The only modification to the data was to correct clearly erroneous route numbers. For example, in some cases the carrier hand-entered the route number, like on a pivot. The carrier would inadvertently enter the letter 'O' instead of the number "0" so the route number would be CO31 instead of C031. Correction of the route numbers in no way affected the scan data. In addition, there are two

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

created variables in that data set. They are named Barcode 1 and Barcode2, and they are both just the last 3 characters of the raw variable named TRACK\_BARCODE\_ID. TRACK\_BARCODE\_ID is the variable that contains the scans of the barcodes taken in the field. Barcode1 and 2 were created because, as explained in the Report at page 92-93, the last 3 characters are enough to uniquely identify the scanned activity. Each observation in the raw data set, Carrier\_scans\_pa\_study.sas7bdat, is an individual scan. Scanrouteday\_masked\_zips.sas7bdat cumulates the individual scans into route day observations, and was developed from the scan data.

Package\_study\_volume\_masked\_zips.sas7bdat contains the raw data volume data collected through the WEBTOOL. All variables except ZIP and ROUTE were downloaded from the WEBTOOL. The variables ZIP and ROUTE were created in SAS by extracting the appropriate characters from the variable ZIP\_ROUTE. The description of other variables from this dataset is contained in the response to question 5a.

- b. The log file that created scanrouteday\_masked\_zips.sas7bdat is filed in RM2014-7/NP1. The variables that were used are listed below, along with a description of each variable.

DATE – date of scan

ZIP – ZIP Code for the route

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

ROUTE\_ID – identifier for the route as recorded on the scan record

TRDV\_DEV\_ID – identifier for the scanner as recorded on the scan record

CURRENT\_TIME – time of the scan

BARCODE1&2 – they both are last three characters from the

TRACK\_BARCODE\_ID As mentioned in the answer to part a., just the last three characters are needed to identify a delivery activity, such as Begin Deviation Parcel Delivery.

There were three steps required to construct the route-day dataset for scan times. They are:

1. Sort the scans into the correct order based on the DATE, ZIP, TRDV\_DEV\_ID, ROUTE\_ID, and CURRENT\_TIME,
2. Calculate the elapsed time between each scanned activity pair, which consists of a 'Begin' activity and ends with a delivery mode scan, and
3. Sum the elapsed time and the number of scan pairs for each delivery activity (like "deliver an accountable") by DATE, ZIP, and ROUTE\_ID.

The log file that created package\_study\_volume\_masked\_zips.sas7bdat is contained in USPS-RM2015-7/1 within the folder

Cost\_Pool\_Formation/SAS\_Logs. The log file is entitled,

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- Package\_Study\_WEBTOOL\_Readin.log. The only modifications made to the raw data were to create the variables ZIP and ROUTE from the raw variable ZIP\_ROUTE by extracting the appropriate characters, and to correct for a minor technical glitch that occurred with the WEBTOOL on March 29, 2014.
- c. Carriers used their regular barcode scanners when recording the special barcodes that were used in the study. As a result, some extraneous barcode data were included in the raw data. EVENT\_CODE, EVENT\_DATE1, and EVENT\_TIME1 are all extraneous variables that were not part of the study, so no definitions were obtained for these variables. TRACK\_BARCODE\_ID, TRDV\_DEV\_ID, Barcode2, and Barcode1 are defined in answer b.
- d. The variable EVENT\_CODE was recorded by the carrier's scanner, but was extraneous to the study and plays no role in it. As a result, the individual event code numbers were not reviewed and definitions for those codes were not obtained (and not needed).
- e. Not confirmed. Barcode1 and Barcode2 identify the delivery activity scanned by the carrier, and have nothing to do with the extraneous variable EVENT\_CODE.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 10.** Please refer to the Report, page 22, which describes the variables included in the regular delivery model.
- a. Please confirm that unless a small parcel was cased in the office, it would be considered an in-receptacle parcel for this study. If there were exceptions to this, please explain.
  - b. For motorized routes or route segments, are small parcels that are not cased in the office typically mixed in with large parcels, or are they typically put in a tray, container, or other location in the vehicle separate from large parcels? If not, how are large and in-receptacle parcels placed in the vehicle?
  - c. For routes or route segments in which the carrier walks to the box to deliver the mail, are small parcels that are not cased in the office typically placed in carrier satchels? If not, please explain.

**RESPONSE:**

- a. Given that a small package is defined as one that can be delivered in the mail receptacle, then the statement is confirmed.
- b. Small packages that are not cased in the office are typically put in a tray, container, or other location in the vehicle separate from large parcels
- c. Yes.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 11.** Please refer to USPS-RM2015-7/1, Collection\_Study\_Exhibit\_2 - Instructions for Local study Coordinators, which states that "No Carrier Pickup parcels are included in the study." Please confirm that carrier pickup parcels are not considered to be collected customer parcels. If not, please explain.

**RESPONSE:**

Confirmed

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

12. Please refer to USPS-RM2015-7/1, Report, page 28, which states that a “sample size of 300 ZIP Codes was determined to be the largest sample...” and “the collection study utilized a stratified systematic sample from a frame of 10,720 ZIP Codes.” Table 10 shows six strata, the number of ZIP Codes per strata in the sample, and the Share of delivery time per strata in the sample.
- a. Please add eight columns to Table 10 so that it also shows the following information:
    - i. the number of ZIP Codes in each strata in the sample frame;
    - ii. the number of ZIP Codes in each strata in the population;
    - iii. the share of street time in each strata in the sample frame;
    - iv. the share of street time in each strata in the population;
    - v. the average street time in each strata of the sample;
    - vi. the average street time in each strata of the population;
    - vii. the variance of street time in each strata of the sample; and
    - viii. the variance of street time in each strata of the population.
  - b. Please describe the characteristics of the sampling frame – e.g., the percentage of city carrier routes included, from which database(s) the sampling frame is constructed and how it is constructed, the representativeness of the sampling frame to the population of city carrier routes.

**RESPONSE:**

a. For the purposes of constructing the sample, the population was defined to be the list of city carrier routes used for the ongoing City Carrier Cost System (CCCS) frame. The sample frame was the set of routes included in the most concurrent Form 3999 Data Set that was available when the sample was selected. A description of the CCCS frame can be found in the ACR in USPS-FY14-34, and a description of the Form 3999 Data Set is contained in the Report starting on page nine.

Note that CCCS uses the Address Management System (AMS) as the basis for its frame, which contains information on delivery points and routes but does not

## RESPONSE OF THE UNITED STATES POSTAL SERVICE TO CHAIRMAN'S INFORMATION REQUEST NO. 1

contain any information regarding work hours. Thus, there are no street times available for the population, so parts iv., vi., and viii. of this question could not be completed. The following table contains the requested information for the remaining parts of the question.

### Response to parts i., ii., iii., v., and vii. for CHIR No. 1 Q 12a.

Stratum	Strata Definition	# ZIPs	Stratum Proportion of Street Time	Sample Frame ZIPs	Population ZIPs	% Street Time Sample Frame	Average Street Time <sup>1</sup> Per ZIP Sample Frame	Variance Street Time <sup>1</sup> Sample Frame
<b>Column Q12</b>	<b>Roman Numerals from CHIR 1 Q12</b>			i.	ii.	iii.	v.	vii.
Small Driving	Driving ZIP < 6 Routes	1,209	2%	1,209	1,344	2%	15.25	63.58
Small Walking	Walking ZIP < 6 Routes	2,087	4%	2,087	2,884	4%	15.14	52.93
Medium Driving	Driving ZIP 5 < x < 21 Routes	2,254	20%	2,254	2,249	20%	67.35	590.96
Medium Walking	Walking ZIP 5 < x < 21 Routes	2,787	24%	2,787	2,789	24%	66.33	648.36
Large Driving	Driving ZIP > 20 Routes	882	18%	882	886	18%	153.25	1,349.88
Large Walking	Walking ZIP > 20 Routes	1,501	32%	1,501	1,506	32%	160.03	1,682.80
Total				10,720	11,658	100%		
<sup>1</sup> Street Time is defined as it defined in the Cost and Revenue Analysis Report. Thus, it does not include loading/unloading the carrier vehicle								

b. The sample frame was the set of routes from the most concurrent Form 3999 data set that was available when the sample was selected. This data set is updated regularly throughout time as newer route evaluations become available and routes are added or eliminated. A detailed description of the Form 3999 data set is provided in the Report starting on page nine.

As described in the Report, the Postal Service is using the Form 3999 dataset as the basis for constructing the cost pools. To provide consistency between the

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

estimated variabilities and the cost pools to which they are applied, it was necessary to have a sample frame that included street times by activity. The Form 3999 data contain these street times and was thus chosen as the basis of selecting the sample. Using routes in the Form 3999 dataset as the sample frame provided an excellent snapshot of the city letter route delivery network as it contains over ninety-eight percent of the routes in the population.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 13.** Please refer to USPS-RM2015-7/1, Collection\_Study\_Exhibit\_2 - Instructions for Local Study Coordinators, which states that the sources of collection mail are 1) “[m]ail collected by regular and auxiliary carriers during delivery, [2] l]etter and flat mail accepted by clerks at the retail window; letter and flat mail collected by clerks from boxes and lobby drops; letter and flat mail dropped off by customers at the dock of Function 4 offices; [and 3)] mail collected by city carriers on dedicated collection runs.”
- a. Please identify which of these sources were included in the study performed in Docket No. R2005-1.
  - b. Please identify which, if any, of these sources were previously included in general or express collections.
  - c. Please confirm that mail from the second source (collected by clerks) is not mail collected by city carriers. If confirmed, please explain whether this mail was included in the collection mail variable in the regular city carrier delivery model. If not confirmed, please explain.
  - d. Please reconcile the description of collection mail sources found in USPS-RM2015-7/1, Collection\_Study\_Exhibit\_2 - Instructions for Local study Coordinators, and the sources of collection mail described in USPS-RM2015-7/1, Report, at 30.

**RESPONSE:**

- a. Mail collected by letter carriers during delivery.
- b. Previous analysis of general and express collections was based upon a special study that measured collection volumes handled by special purpose carriers when they swept street letter boxes on collection runs. The analog to that special study among the sources listed in the question would be mail collected by city carriers on dedicated collection runs.
- c. Confirmed. It is not included in the collection mail variable in the regular delivery time city carrier model.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

d. Only mail collected by carriers during delivery is included in the regular delivery time model, as was the case in the study performed for Docket No. R2005-1. The collection mail study was done in conjunction with Delivery Operations, which increased cooperation by field employees and ensured consistency with postal operations. Delivery Operations requested that measurements of the volumes collected from the second two sources be included in the field study for its purposes. These data were collected solely for those Operations purposes and were not, in any way, included in the city carrier street time study. The sources of collection mail described in USPS-RM2015-7/1, Report, at 30 are accurate and reflect what was done in the study. The actual data collection instructions, including definitions of the sources of data used solely by Operations, were submitted in the exhibit.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

14. Please refer to USPS-RM2015-7/1, Report at 34, which states that “there were a sufficient number of reported collection volume days so that it was legitimate to use mean values over the days for which the route reported volume as the imputed values.”
- a. Please provide the mean and median values for each imputed value by route.
  - b. Please discuss the circumstances under which the median value would be a more appropriate imputation value than the mean.
  - c. Please discuss whether any other imputation methods were considered or employed. If any other imputation methods were tested, please provide the data, programs, and logs for these tests. In the response, please include a discussion of the following potential alternatives:
    - i. Simple random imputation, which randomly imputes missing values based on the observed data for the variable with missing values
    - ii. Matching, which, in this case, would impute values using observations with similar delivery characteristics (e.g., uses the same delivery technology and has a similar business delivery ratio).
    - iii. Multiple Imputation (See e.g., the Stata MI procedure).
  - d. Was an alternative analysis performed with a dataset that excludes incomplete records instead of imputing values for them? If so, please provide the data, programs, and logs from this alternative. If not, please detail the impact of using imputed data in the analysis.

**RESPONSE:**

- a. The attached workbook named **CHIR.1.Q14a.Resp.xlsx** contains the mean and median for each imputed value.
- b. The median may be a more appropriate imputation value for routes with highly skewed distributions. However, as one can see from the workbook filed in response to part a, the mean and median collection volumes are typically quite similar.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

c. No other imputation methods were considered. Given that the analysis revolved around imputing volumes for just 206 out of over 72,000 route days, efficient resource allocation dictated that only one imputation method be used. The Postal Service has limited resources for doing cost studies, so efficient use of those resources is important. Given the small number of routes with imputed volumes, it is highly unlikely that different methods of imputation would have a material impact on the results, so no alternative methods were investigated.

d. The impact of imputation is likely to be *de minimis* because so little imputation was done. Moreover, when imputation was done, it was often done for just one route in a ZIP Code, so the imputation likely had relatively small impact on the ZIP Code's daily collection volume. Finally, quite a few of the route days with missing data turned out to be for vacant routes. In those instances, the actual and imputed volumes were both zero, so imputation was "perfect." In sum, imputation was probably not necessary, but it provided the advantage of ensuring that all ZIP Code days used in the analysis were complete.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

15. Please compare the Report at 38, where the Postal Service states that “[t]he mean number of pieces collected per day per ZIP Code was 3,520,” and Table 14 (on page 39), which shows the mean collection volume as 173.6 pieces.
- a. Please reconcile these two measures.
  - b. Please provide a version of Table 14 using collection volumes per ZIP code.

**RESPONSE:**

a. The measure of 3,520 pieces refers to collection volumes per ZIP Code day and the measure of 173.6 pieces refers to collection volumes per route day.

Table 14 presents result in terms of volume per route because the City Carrier Cost System does not produce volumes per ZIP Code.

b. While there is no direct measure of volume per ZIP Code from the City Carrier Cost System, a measure of ZIP code volume can be produce by multiplying the route day values by the average number of routes per ZIP code. That is what was done to produce the following table:

Comparing Measures of Central Tendency for Two  
Different Sets of Collection Data at the ZIP Code Level

Source	Mean	Median
Collection Volume Study	3,520.0	1,918.2
FY 2012 City CCS	3,678.2	1,926.3

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 16.** Please refer to the Report at pages 43-44, which states that “[o]n any given day, the actual allied time may differ from the systematic allied time because of random factors, such as more or less traffic on the route, or a variation in a carrier’s personal needs time.”
- a. Please confirm that allied time is calculated from Form 3999, or route evaluation data from the sample period Form 3999 data. If not, please explain.
  - b. Please identify the source of allied time in Table 16 (on page 42). Please also identify:
    - i. The number of ziproute days used to calculate allied time;
    - ii. The date ranges used calculate allied time, by activity if they are not the same; and
    - iii. Whether or not these allied time values are limited to the study period and route days used in the regular delivery model before or after observation reduction.

**RESPONSE:**

- a. Confirmed
- b.i. Table 16 represents the average allied time over 112,972 route days.
- b.ii. The same 112,972 route days are used for all activities.
- b.iii. The allied time proportions in Table 16 are based upon on all route evaluations done in FY 2012 and FY 2013. The allied times used in the regular delivery model are just for the routes included in the data set used to estimate the regular delivery time equation, which is a smaller number. Please note that the proportions presented in Table 16 are purely for illustration and are not used in estimating the regular delivery equation.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 19.** Please refer to USPS-RM2015-7/1, In\_Receptacle\_Variabilities\_Model.SAS  
Please confirm that in-receptacle parcels that were accountables were only counted as accountables, and were not include in the model. If not confirmed, please explain.

**RESPONSE:**

Confirmed. Any package that was also an accountable was counted as an accountable. Because such a package was not delivered in the receptacle, it was not included in the in-receptacle delivery equation.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 20.** Please refer to the Report at page 87, where the Postal Service states that “the package delivery time equations will include the proportions of the package and accountable deliveries made by mode.”
- a. Please confirm the rationale for this decision is provided on the same page, by the statement “[b]ecause there are so few package and accountable deliveries made on a route, it could well be that the route’s overall delivery type does not reflect the delivery technology for its package and accountable delivery.” If not, please explain.
  - b. Please explain what investigation was made to confirm the statement in part a. Please provide all data, programs, and logs used in the investigation.
  - c. Did the Postal Service run a regression using delivery technologies, rather than delivery mode, in a package delivery model? If so, please provide all data, programs, and logs necessary to reproduce the results.

**RESPONSE:**

- a. Confirmed
- b. This investigation involved comparing the average number of packages delivered on a route, about 40, with the typical number of delivery points on a route, about 600. Because some delivery points receive more than one package at a time, the 40 packages imply that fewer than 40 delivery points receive packages. If, for example, the number of delivery points receiving packages is 30, then just 5 percent of delivery points on a route receive packages. It is a reasonable inference that the nature of package delivery is more aptly described by the delivery types for the 5 percent of stops that get packages than by the other 95 percent that do not.
- c. No.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 21.** Please refer to the Report at page 99 where the Postal Service states that the "average time per study scan ... [was] 12 seconds." Please explain whether average scan time was the average of the elapsed time between all before and after delivery scans. If this was not the case, please explain how scan time for each type of delivery was measured.

**RESPONSE:**

Not confirmed. First, please note that the discussion on page 99 of the Report refers to the average time it took to scan a barcode during the study, not the average elapsed time for the delivery activities being studied. That latter average would be measured by the average of the elapsed time between begin activity scan and the mode scan (which was also the end activity scan). The average time for scanning a barcode was found by asking participating ZIP Codes to keep track of the additional daily street time that was required to complete the study scans. This would be the time spent by carriers in scanning. Dividing scanning time by the associated number of barcode scans made yielded the average time per scan.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 22.** Please refer to the Report at pages 99-100 where there is a discussion of the problem of zero-second scans due to the carrier's failure to "hit the 'A' key for the begin activity scan."
- a. Please provide the raw data used to correct for this problem.
  - b. Please describe each variable in the raw data, and provide all data, programs and logs needed to make the adjustments to correct for this problem.

**RESPONSE:**

a. This correction is made in the SAS programs that estimate the package and accountable delivery time variability equations. Both of these programs are provided in RM2015-7/1. For example, the correction to the data used to estimate the in-receptacle package equation is made in the SAS program entitled, In\_Receptacle\_Variabilities\_Model which is in the folder entitled "In Receptacle Parcel Equation\SAS Programs." (Note that the same correction is made in both programs.) The data used in that program are also presented in RM2015-7/1 in the folder entitled: "In Receptacle Parcel Equation\SAS Data Sets."

b. The SAS code that makes the correction is presented in the SAS programs used to estimate the package and accountable delivery time models. It is repeated below for convenience. Note that the SAS code implements the following computational formula, which is provided on page 100 of the report:

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

$$\text{Adjusted Time} = \sum_{i=1}^m \frac{\text{Net Delivery}}{\text{Time}} \Bigg|_{\text{Time} > 0} * \frac{(\text{Total \# of Deliveries})}{(\text{\# of Deliveries With Positive Time})}$$

The relevant SAS Code is:

```
if DevAcctCT > 0 then DevAcctT = DevAcctT * (DevAcctC/DevAcctCT);  
else DevAcctT=0;  
  
if DevPackCT > 0 then DevPackT = DevPackT * (DevPackC/ DevPackCT);  
else DevPackT=0;  
  
if InReceptCT > 0 then InReceptT = InReceptT * (InReceptC/  
InReceptCT); else InReceptT=0;  
  
if MoveVehCT > 0 then MoveVehT = MoveVehT * (MoveVehC/ MoveVehCT);  
else MoveVehT=0;
```

The variable definitions are presented below:

DevAcctCT = Number of accountable deliveries with positive time.

DevAcctT = Accountable delivery time

DevAccC = Total number of accountable deliveries

DevPackCT = Number of deviation package deliveries with positive time.

DevPackT = Deviation package delivery time

DevPackC = Total number of deviation package deliveries

InReceptCT = Number of in-receptacle package deliveries with positive time.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

InReceptT = In-receptacle package delivery time

InReceptC = Total number of in-receptacle package deliveries

MoveVehCT = Number of move vehicle deliveries with positive time.

MoveVehT = Move vehicle delivery time

MoveVehC = Total move vehicle deliveries

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 23.** Please explain how the time inserting an in-receptacle parcel was separated from the time inserting mail (which was counted in the regular delivery model) jointly inserted into the same receptacle.
- a. If the time was not separated, please explain.
  - b. If the time was separated, please provide the raw data and SAS or other programs and logs needed to identify scan and insertion time associated with this type of package delivery and to separate it from regular mail delivered at the same time to the same receptacle.

**RESPONSE:**

Although in-receptacle packages, for a given receptacle, are delivered at the same stop at which the letters and flats for that receptacle are delivered, they are not delivered in the same motion. This is because, in part, of the need to scan the product tracking barcodes on the parcels. To ensure accurate capture of the separate time required for delivering in-receptacle packages, carriers were instructed to first deliver the letters and flats into the receptacle and then to subsequently deliver any in-receptacle packages. Immediately preceding the initiation of the in-receptacle package delivery, the carrier would scan the Begin In-receptacle Delivery barcode. He or she would then deliver the package and, once the delivery was over, scan the Mode barcode, which also served as the end delivery barcode. The in-receptacle delivery time was separated from the regular delivery time in the collected data by identifying the barcode scans associated with in-receptacle package delivery and measuring the time that elapsed between the begin delivery scan and the mode scan.

- a. Not applicable.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

b. The raw data used to identify and separate in-receptacle package delivery time is in the SAS data set entitled, carrier\_scans\_pa\_study.sas7bdat, which is in RM2015-7/NP1. The program that calculates and separates the in-receptacle package time is the SAS program entitled "Scan.Route.Day." It is also provide in RM2015-7/NP1.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

24. Please refer to the Report at pages 101-103, where the Postal Service discusses methods to resolve several data problems that were identified in the parcel/accountable and in-receptacle measurement of delivery time. Please provide all raw data, explain all variables, and provide all programs and logs used to make the data improvements discussed.

**RESPONSE:**

The data improvements discussed on pages 101-103 of the report were made in SAS programs that estimate the package and accountable delivery time variability equations. The correction to the data used to estimate the in-receptacle package equation is made in the SAS program entitled, In\_Receptacle\_Variabilities\_Model which is in the folder entitled "In Receptacle Parcel Equation\SAS Programs". The correction to the data used to estimate the deviation package and accountable equation is made in the SAS program entitled, Deviation\_Acct\_Variabilities\_Model which is in the folder entitled "In Deviation Parcel Acct Equation\SAS Programs".

For data used to estimate the in-receptacle package equation, the improvements were made with the following SAS code:

```
data analysis_three; set analysis_three;
if irpv > 0 then min_per_irpc =(irpt/irpv)/60; else min_per_irpc=0;
data analysis_three; set analysis_three;
route_gap=nvroutes-nsroutes;
pct_route_gap=route_gap/does_routes;
```

## RESPONSE OF THE UNITED STATES POSTAL SERVICE TO CHAIRMAN'S INFORMATION REQUEST NO. 1

```
ir_pc_pd=irpv/irpd;  
  
if min_per_irpc > 3 then delete;  
  
if min_per_irpc < 0.166 and pct_route_gap gt 0.25 then delete;  
  
if min_per_irpc < 0.166 and ir_pc_pd lt 2  
then delete;
```

For data used to estimate the deviation package and accountable equation, the improvements were made with the following SAS code:

```
data analysis_three; set analysis_three;  
  
route_gap=nvroutes-nsroutes;  
  
pct_route_gap=route_gap/does_routes;  
  
if min_per_devpc gt 5 then delete;  
  
if min_per_devac gt 10 then delete;  
  
move_ratio=mvt/devtime;  
  
if min_per_devpc < 0.166 and pct_route_gap gt 0.25 or  
min_per_devpc < 0.166 and move_ratio lt .25 then delete;  
  
if min_per_devac > 0 and min_per_devac < 0.3333 and pct_route_gap gt  
0.25 then delete;  
  
if min_per_devac > 0 and min_per_devac < 0.3333 and move_ratio lt .25  
then delete;
```

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 25.** Please refer to the Report at page 103, where the Postal Service states that in response to observing many ZIP Code days with large package and accountable time, the Postal Service performed “[f]urther investigation of these large ZIP Code days [which] suggested that their cost generating process may be different from the rest of the ZIP Code days.”
- a. Please describe the further investigation to which this quote refers.
  - b. If the investigation involved quantitative analysis, please provide all data and programs needed to reproduce the results.
  - c. Please also provide all data and programs needed to reproduce the investigation referred to on pages 105-109.

**RESPONSE:**

- a. The investigation involved separately examining the relationship between delivery time and delivery volume for the ZIP Code days with more than 8 hours of package delivery time, and the ZIP Code days with less than 8 hours of delivery time. The investigation involved creating cross plots for these two variables for each subset and examining if the patterns in the cross plots were the same or different.
- b. Not applicable.
- c. The investigation described on pages 105 to 109 involved looking at the means for the characteristic variables included in the analysis data set across the two data subsets, to see if there were any obvious differences. The required data and programs are presented in RM2015-7/1. The required data are the analysis data used to estimate the package/accountable delivery time equations. The required programs are the SAS programs called In-Receptacle\_Variabilities\_Model and Deviation\_Acct\_Variabilities\_Model. For

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

convenience, the ZIP Code day averages for the characteristic variables are presented below.

Average Values for Characteristic Variables for In-Receptacle Delivery

<b>Variable</b>	<b>High Delivery Time ZIP Code Days</b>	<b>Normal Delivery Time ZIP Code Days</b>
% Door Deliveries	55.4%	42.3%
% Curb Deliveries	16.4%	32.3%
% Dismount Deliveries	10.1%	8.3%
% Central Deliveries	7.5%	5.8%
% Clusterbox Deliveries	10.6%	11.3%
Miles Per Deliver Point	0.001	0.010
% Business Deliveries	9.1%	9.0%

Average Values for Characteristic Variables for Deviation Delivery

<b>Variable</b>	<b>High Delivery Time ZIP Code Days</b>	<b>Normal Delivery Time ZIP Code Days</b>
% Door Deliveries	43.8%	23.6%
% Curb Deliveries	16.2%	16.4%
% Dismount Deliveries	23.7%	17.8%
% Central Deliveries	7.3%	8.7%
% Clusterbox Deliveries	9.1%	9.7%
Miles Per Deliver Point	0.001	0.004
% Business Deliveries	8.8%	4.3%

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 26.** Please refer to the Report at pages 108-109 where it discusses the Grid Search Method employed to develop the optimum threshold value for the in-receptacle switching equation.
- a. Please provide all programs and logs used to derive the optimum threshold for the in-receptacle equation.
  - b. Did the Postal Service use a cross validation method to confirm the significance of the estimated threshold value? If not please explain.
  - c. If a cross validation method was employed, please all raw data, explain all variables, provide all programs and logs used to cross validate.

**RESPONSE:**

a. The program used to do the grid search is provided in RM2015-7/1. It is the program entitled, In\_Receptacle\_Variabilities\_Model.sas in the folder entitled "In Receptacle Parcel Equation\SAS Programs." The search for the optimal threshold value was performed through a coarse-to-fine grid search evaluation scheme. In this type of grid search, the threshold is initially identified through a large step grid in order to identify the broad area in which the optimum occurs. Then, a subsequent, denser, grid search is performed that focuses just on the broad area identified in the initial search. The process is repeated until the optimum is found. Note that as the area becomes smaller, the grid step also becomes smaller and the search becomes finer.

Mechanically the grid search was done by re-running the SAS program entitled, In\_Receptacle\_Variabilities\_Model, with different values for the threshold value and recording the resulting Root Mean Square Error values. Because the same program was run repeatedly, the logs were not retained. However, the values for

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

the RMSE associated with each grid search were retained and are presented below.

The first grid search covered the range from 22,000 seconds (6.1 hours) to 35,000 seconds (9.7) hours with grid steps of 1,000 seconds. The resulting RSME values are presented next:

Initial Grid Search for IR Package Threshold

Threshold Value	RMSE
22,000	5,848.22
23,000	5,760.49
24,000	5,673.12
25,000	5,641.29
26,000	5,590.68
27,000	5,575.94
28,000	5,557.16
29,000	5,538.02
30,000	5,556.16
31,000	5,581.45
32,000	5,582.79
33,000	5,605.57
34,000	5,681.88
35,000	5,725.87

The smallest RMSE errors occur in the range between 27,000 and 30,000 seconds, so that becomes the range for the second search. In this second grid

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

search, the grid steps are much finer, at a value of just 100 seconds. The results of the second grid search are presented in the following table:

Second Grid Search for IR Package Threshold

Threshold Value	RMSE
27,000	5,575.94
27,100	5,570.12
27,200	5,573.03
27,300	5,566.34
27,400	5,569.85
27,500	5,568.19
27,600	5,572.53
27,700	5,568.40
27,800	5,561.91
27,900	5,560.74
28,000	5,557.16
28,100	5,547.60
28,200	5,542.72
28,300	5,542.72
28,400	5,542.89
28,500	5,544.88
28,600	5,542.63
28,700	5,543.49
28,800	5,541.31
28,900	5,543.58
29,000	5,538.02
29,100	5,538.08

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

29,200	5,535.55
29,300	5,533.90
29,400	5,533.90
29,500	5,533.90
29,600	5,536.96
29,700	5,542.32
29,800	5,555.46
29,900	5,557.81
30,000	5,556.16

The second grid search produces a minimum RMSE for three values, 29,300 seconds, 29,400 seconds, and 29,500 seconds. This suggests that there are no data points between 29,300 and 29,500. This was confirmed with a final grid search between 29,300 and 29,500 with a grid step of 25.

**Third Grid Search for IR Package Threshold**

<b>Threshold Value</b>	<b>RMSE</b>
29,300	5,533.90
29,325	5,533.90
29,350	5,533.90
29,375	5,533.90
29,400	5,533.90
29,425	5,533.90
29,450	5,533.90
29,475	5,533.90
29,500	5,533.90

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

b. No. Given the expense and effort in collecting the field data necessary for estimating the package and accountable delivery models, it was infeasible to collect a sufficiently large amount of data to provide for a reliable cross validation.

c. Not applicable

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 27.** Please refer to the Report at pages 113 where it mentions the Postal Service used the same Grid Search Method employed to develop the optimum threshold value for the equation for the deviation delivery switching equation.
- a. Please provide all programs and logs used to derive the optimum threshold for the deviation delivery switching equation.
  - b. Did the Postal Service use a cross validation method to confirm the significance of the estimated threshold value? If not, please explain.
  - c. If a cross validation method was employed, please all raw data, explain all variables, provide all programs and logs used to cross validate.

**RESPONSE:**

a. The program used to do the grid search is provided in RM2015-7/1. It is the program entitled, Devatiation\_Acct\_Variabilities\_Model.sas in the folder entitled "Deviation Parcel Acct Equation\SAS Programs." The search for the optimal threshold value was performed through a coarse-to-fine grid search evaluation scheme. In this type of grid search, the threshold is initially identified through a large step grid in order to identify the broad area in which the optimum occurs. Then, a subsequent, denser, grid search is performed that focuses just on the broad area identified in the initial search. The process is repeated until the optimum is found. Note that as the area becomes smaller, the grid step also becomes smaller and the search becomes finer.

Mechanically the grid search was done by re-running the SAS program entitled, Devatiation\_Acct\_Variabilities\_Model, with different values for the threshold value and recording the resulting Root Mean Square Error values. Because the same program was run repeatedly, the logs were not retained. However, the

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

values for the RMSE associated with each grid search were retained and are presented below.

The first grid search covered the range from 10,000 seconds (2.8 hours) to 35,000 seconds (6.7) hours with grid steps of 1,000 seconds. The resulting RSME values are presented next:

Initial Grid Search for Deviation Package  
Threshold

Threshold Value	RMSE
10,000	6,468.58
11,000	6,310.87
12,000	6,242.94
13,000	6,170.05
14,000	6,026.29
15,000	6,042.36
16,000	6,087.73
17,000	6,036.17
18,000	6,072.70
19,000	6,078.02
20,000	6,089.62
21,000	6,080.81
22,000	6,125.05
23,000	6,181.23
24,000	6,233.26

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

The first grid search found 14,000 to be the preliminary threshold value with the lowest RMSE, so the second grid search range will be centered on that value. The second grid search range was thus from 13,500 seconds to 15,000 seconds with a grid step of 100. The results of this finer grid search are presented next:

Second Grid Search for Deviation Package  
Threshold

Threshold Value	RMSE
13,500	6,067.04
13,600	6,043.42
13,700	6,033.55
13,800	6,038.75
13,900	6,010.87
14,000	6,026.29
14,100	6,007.48
14,200	6,020.05
14,300	6,022.73
14,400	6,019.58
14,500	6,015.27
14,600	6,017.88
14,700	6,035.03
14,800	6,047.16
14,900	6,043.18
15,000	6,042.36

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

The second grid search found 14,100 to be the preliminary threshold value with the lowest RMSE, so the third grid search range will be centered on that value. In addition, as the range is constricted to the area around that value (from 14,000 to 14,200) the grid step is made smaller at a value of 25. The results of the third grid search are given next:

Third Grid Search for Deviation Package  
Threshold

Threshold Value	RMSE
14,000	6,026.29
14,025	6,023.53
14,050	6,007.54
14,075	6,010.38
14,100	6,007.48
14,125	6,014.65
14,150	6,015.11
14,175	6,017.13
14,200	6,020.05

The third grid search also provides a threshold value of 14,100. The final grid search is thus also centered on this value. The fourth grid step is smaller at a value of only 10 seconds and it covers a smaller range around 14,100 from 14,070 to 14,300. It produces the final threshold value of 14,080:

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

Fourth Grid Search for Deviation Package

Threshold

Threshold Value	RMSE
14,070	6,010.38
14,080	6,005.91
14,090	6,007.48
14,100	6,007.48
14,110	6,007.69
14,120	6,014.65
14,130	6,015.63

b. No. Given the expense and effort in collecting the field data necessary for estimating the package and accountable delivery models, it was infeasible to collect a sufficiently large amount of data to provide for a reliable cross validation.

c. Not applicable

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

- 28.** This question seeks further information on the calculation of in-receptacle delivery time and its relation to regular delivery time.
- a. Please confirm in-receptacle packages are delivered with the customer's "regular delivery mail." If not, please explain.
  - b. Was all in-receptacle scan time subtracted from regular delivery street time? If not, please explain.
  - c. Please provide all raw data, explain all variables, and provide all programs, quantitative analyses, and logs used to calculate the time associated with the placement of an in-receptacle parcel in a customer's "box." If this was not calculated, please explain why not.

**RESPONSE:**

a. Both regular delivery time and package delivery time are part of total street time. However, package delivery time is part of allied time, not regular delivery time. In-receptacle packages were delivered at the same stop as the associated letters and flats, but the in-receptacle packages were placed separately into the receptacles that received the letters and flats. This was because, in part, of the need to scan the product tracking barcodes on the parcels. Carriers first delivered the letters and flat to the receptacle and then separately delivered the in-receptacle packages. As a result, the time for delivering the in-receptacle packages was measured separately from any time spent delivering letters and/or flats.

b. Yes. As stated on pages 17 and 19 of the report:

Because the route evaluation process incorporates in-receptacle package delivery time into regular delivery time, the regular delivery time proportion is overstated for attributable costing purposes. Accuracy requires using the independently measured in-receptacle package delivery time

**RESPONSE OF THE UNITED STATES POSTAL SERVICE  
TO CHAIRMAN'S INFORMATION REQUEST NO. 1**

proportion to reduce the route evaluation delivery time proportion. This calculation is done in Table 8.

c. Please see the SAS program entitled Cost\_Pools\_Parcel\_Acct\_Time in RM2015-7/1 in the folder entitled "Cost Pool Formation" for the calculation of the time associated delivering in-receptacle packages.